บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ต้องการศึกษาและเปรียบเทียบวิธีการทดสอบความสมมูลของตัวแปรสุ่มทวินามภายใต้แนวความคิด 2 แนวคิด คือ แนวคิดแบบคลาสสิก และแนวคิดแบบเบส์ ด้วยการจำลองค่าสัดส่วนการปฏิเสธสมมติฐาน H_0 เมื่อ H_0 เป็นจริง และการจำลองค่าสัดส่วนการปฏิเสธสมมติฐาน H_0 เมื่อ H_0 เป็นเท็จ ภายใต้สถานการณ์ที่แตกต่างกันทั้งสิ้น 108 สถานการณ์ ดังที่ได้กล่าวมาแล้วในบทที่ 2 ดังนั้นในบทนี้จะกล่าวถึง ขั้นตอนดำเนินการวิจัย ซึ่งแบ่งเป็นหลายขั้นตอนย่อยดังต่อไปนี้

ขั้นตอนดำเนินการวิจัย

1. สร้างการทดสอบสมมูลของตัวแปรสุ่มทวินามภายใต้แนวคิดแบบเบส์

2. สุ่มตัวอย่างที่แจกแจกแบบทวินามในแต่ละสถานการณ์ 1000 ครั้ง

3. จำลองค่าสัดส่วนการปฏิเสธสมมติฐาน H_0 เมื่อ H_0 เป็นจริง และ H_0 เป็นเท็จ สำหรับการทดสอบสมมูลภายใต้แนวคิดแบบคลาสสิก โดยใช้วิธีประยุกต์ของพาเทิล และกุปตา (Modified Patel-Gupta test)

4. เตรียมข้อมูลสำหรับการทดสอบสมมูลภายใต้แนวคิดแบบเบส์

5. จำลองค่าสัดส่วนการปฏิเสธสมมติฐาน H_0 เมื่อ H_0 เป็นจริงและ H_0 เป็นเท็จ สำหรับการทดสอบสมมูลภายใต้แนวคิดแบบเบส์

6. สรุปผลการเปรียบเทียบระหว่างสองวิธีในสถานการณ์ต่างๆ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รายละเอียดของแต่ละขั้นตอนเป็นดังนี้

1. สร้างการทดสอบสมมูลของตัวแปรสุ่มทวินามภายใต้แนวคิดแบบเบส์

การสร้างการทดสอบสมมูลภายใต้แนวคิดแบบเบส์นั้น มีพื้นฐานจากตัวประมาณเบส์ นั้นคือต้องทราบลักษณะการแจกแจงก่อน (informative prior distribution) หรือการแจกแจงของพารามิเตอร์ θ ของตัวแปรสุ่มทวินาม เมื่อ $X \sim B(n,\theta)$ แต่ถ้าไม่ทราบลักษณะการแจกแจง

ก่อน (noninformative prior distribution) จะกำหนดให้พารามิเตอร์ θ มีโอกาสเป็นได้ทุกค่าในช่วงที่เป็นไปได้เท่าๆกัน นั่นคือกำหนดให้ θ_1 และ θ_2 มีการแจกแจงเหมือนกันคือ $Uniform(0,1)$ กล่าวคือ $\pi(\theta_1)=\pi(\theta_2)=1$ ซึ่งในงานวิจัยครั้งนี้ได้กำหนดขอบเขตการศึกษาในกรณีไม่ทราบลักษณะการแจกแจงก่อน

เนื่องจาก $\theta \sim Uniform(0,1)$ และ $\theta \sim Beta(\alpha=1,\beta=1)$ คือการแจกแจงแบบเดียวกัน ดังนั้นในขั้นตอนต่อไปจะพิจารณาที่ θ ที่มีการแจกแจงแบบ $Beta(\alpha,\beta)$ และเมื่อแล้วเสร็จจึงแทนค่า $\alpha=1,\beta=1$ กลับคืน เพื่อความสะดวกในกรณีที่ทราบลักษณะการแจกแจงก่อนจะสามารถแทนค่าพารามิเตอร์ α,β เพื่อกำหนดโค้งการแจกแจงก่อนได้

กำหนดการแจกแจงของ X เมื่อกำหนด θ (sampling distribution) จาก $X \sim Bi(n,\theta)$ นั่นคือ $f(x|\theta)=\binom{n}{x}\theta^{x}(1-\theta)^{n-x}$

การหาการแจกแจงส่วนริม (marginal distribution: $m(x)$) ของ X จากการแจกแจงร่วม (joint distribution : $f(x,\theta)$) ของ X และ θ โดย

$$m(x)=\int_0^1 f(x,\theta)d\theta \qquad \text{เนื่องจาก } 0\le\theta\le 1$$

$$=\int_0^1 f(x|\theta)\pi(\theta)d\theta$$

$$=\int_0^1 \binom{n}{x}\theta^{x}(1-\theta)^{n-x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}d\theta$$

$$=\binom{n}{x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 \theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}d\theta$$

$$=\binom{n}{x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)}{\Gamma(n+\alpha+\beta)}$$

การหาการแจกแจงหลัง (posterior distribution) ซึ่งเป็นการแจกแจงที่มีเงื่อนไขของ θ เมื่อกำหนดตัวอย่าง X ดังนี้

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{m(x)}$$

$$= \frac{\binom{n}{x}\theta^x(1-\theta)^{n-x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\binom{n}{x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)}{\Gamma(n+\alpha+\beta)}}$$

$$= \frac{\Gamma(n+\alpha+\beta)}{\Gamma(x+\alpha)\Gamma(n-x+\beta)}\theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}$$

จะเห็นว่า θ มีการแจกแจงแบบ $Beta(x+\alpha, n-x+\beta)$ นั่นเอง ซึ่ง θ_1 และ θ_2 มีการแจกแจงแบบเดียวกันและเป็นอิสระต่อกัน (พิจารณา θ เป็นตัวแปรสุ่ม และ x เป็นค่าคงที่ซึ่งได้จากการสุ่มตัวอย่าง)

สำหรับกรณีที่ไม่ทราบลักษณะการแจกแจงก่อน จะกำหนดให้พารามิเตอร์ θ มีโอกาสเกิดขึ้นได้เท่าๆกัน (noninformative prior distribution) หรือนั่นคือ $\theta \sim Uniform(0,1)$ หรือ $\theta \sim Beta(\alpha=1, \beta=1)$ ดังนั้น การแจกแจงหลัง (posterior distribution) ในกรณีนี้คือ

$$\pi(\theta|x) = \frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)}\theta^x(1-\theta)^{n-x}$$

ในการอนุมานเกี่ยวกับ θ โดยวิธีเบส์นั้น แตกต่างจากวิธีคลาสสิกคือ จะใช้การแจกแจงหลัง $\pi(\theta|x)$ แทนการแจกแจงของ $f(x|\theta)$ ในวิธีคลาสสิก และการทดสอบสมมติฐานโดยวิธีเบส์นั้นสามารถทำได้โดย ถ้า $P(\theta \in \Theta_0^c|X) > k$ จะปฏิเสธสมมติฐาน H_0 เมื่อ k คือค่าคงที่ที่มีค่ามากจำนวนหนึ่งเช่น 0.99 เป็นต้น สำหรับสมมติฐานในกรณีนี้คือ $H_0 : |\theta_1 - \theta_2| \geq \Delta$ และ $H_1 : |\theta_1 - \theta_2| < \Delta$ ดังนั้น

$$P(\theta \in \Theta_0^c|X) = P(|\theta_1 - \theta_2| < \Delta|X)$$

$$= P(-\Delta < \theta_1 - \theta_2 < \Delta)$$

$$= \iint_R f(\theta_1, \theta_2)dR$$

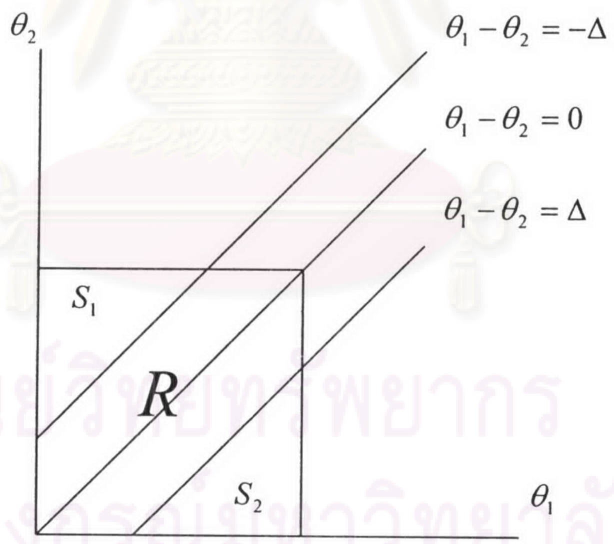
เมื่อ $f(\theta_1,\theta_2)$ คือการแจกแจงร่วม(joint distribution) ระหว่าง θ_1 และ θ_2 ซึ่งได้จาก $f(\theta_1,\theta_2)=\pi(\theta_1|x_1)\cdot\pi(\theta_2|x_2)$ เพราะว่า θ_1 และ θ_2 เป็นอิสระต่อกัน และมีการแจกแจงเหมือนกัน นั่นคือ $Beta(x_1+\alpha_1,n_1-x_1+\beta_1)$ และ $Beta(x_2+\alpha_2,n_2-x_2+\beta_2)$ ตามลำดับ ดังนั้น

$$f(\theta_1,\theta_2)=\frac{\theta_1^{x_1+\alpha_1-1}(1-\theta_1)^{n_1-x_1+\beta_1-1}}{B(x_1+\alpha_1-1,n_1-x_1+\beta_1)}\cdot\frac{\theta_2^{x_2+\alpha_2-1}(1-\theta_2)^{n_2-x_2+\beta_2-1}}{B(x_2+\alpha_2-1,n_2-x_2+\beta_2)}$$

เมื่อ
$$B(x_1+\alpha_1,n_1-x_1+\beta_1)=\frac{\Gamma(x_1+\alpha_1)\Gamma(n_1-x_1+\beta_1)}{\Gamma(x_1+\alpha_1+n_1-x_1+\beta_1)}$$

และ
$$B(x_2+\alpha_2,n_2-x_2+\beta_2)=\frac{\Gamma(x_2+\alpha_2)\Gamma(n_2-x_2+\beta_2)}{\Gamma(x_2+\alpha_2+n_2-x_2+\beta_2)}$$

ซึ่ง $f(\theta_1,\theta_2)$ เป็นฟังก์ชันความน่าจะเป็น ดังนั้น $\int_0^1\int_0^1 f(\theta_1,\theta_2)d\theta_1 d\theta_2=1$ และบริเวณ R คือโดเมนของการอินทิเกรต ดังรูป



รูปที่ 3.1

$$\iint_R f(\theta_1,\theta_2)dR=1-\iint_{S_1} f(\theta_1,\theta_2)dS_1-\iint_{S_2} f(\theta_1,\theta_2)dS_2$$

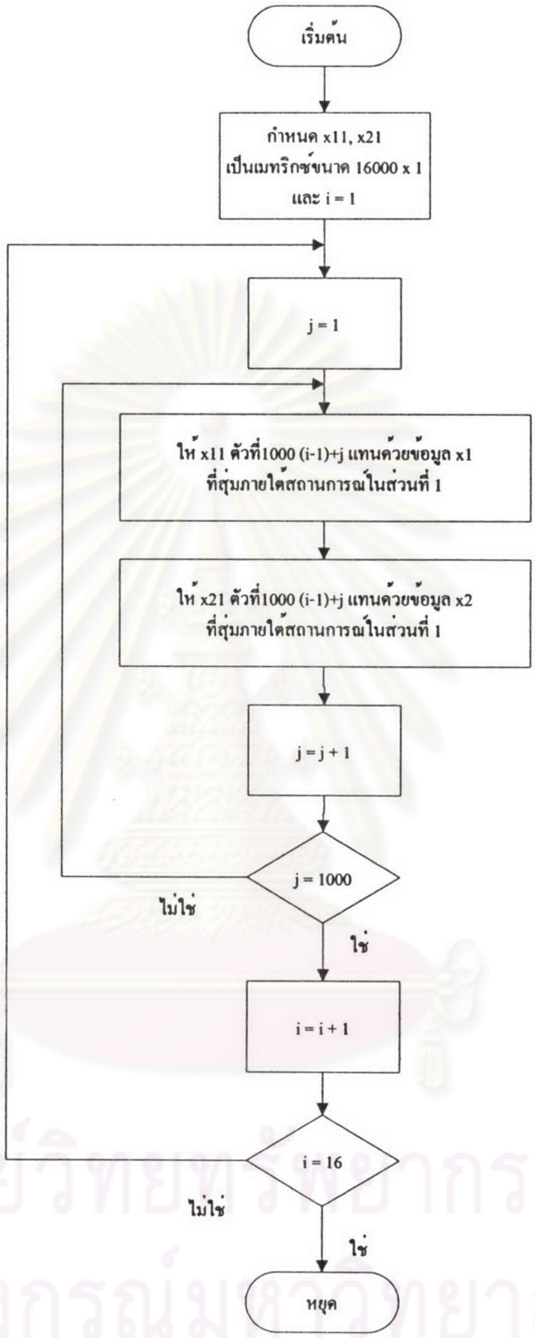
$$=1-\int_{\Delta}^{1}\int_{0}^{\theta_2-\Delta} f(\theta_1,\theta_2)d\theta_1 d\theta_2-\int_{0}^{1-\Delta}\int_{\theta_2+\Delta}^{1} f(\theta_1,\theta_2)d\theta_1 d\theta_2$$

ดังนั้นจะปฏิเสธสมมติฐาน H_0 เมื่อ $\iint_R f(\theta_1,\theta_2)dR > k$ แต่เนื่องจากฟังก์ชัน $f(\theta_1,\theta_2)$ เป็นฟังก์ชันที่ไม่สามารถอินทิเกรตได้ด้วยวิธีทางแคลคูลัส ดังนั้นจะใช้วิธีในการประมาณค่าด้วยโปรแกรมสำเร็จรูปทางคณิตศาสตร์ และจะกล่าวถึงในขั้นตอนต่อไป

2. สุ่มตัวอย่างที่แจกแจกแบบทวินามในแต่ละสถานการณ์ 1000 ครั้ง

กำหนดให้ $X_1 \sim B(n_1,\theta_1)$ และ $X_2 \sim B(n_2,\theta_2)$ ทำการสุ่มตัวอย่าง X_1 และ X_2 ด้วยการกำหนดค่า $\theta_1,\theta_2,\Delta,n_1$ และ n_2 144 สถานการณ์ เป็นจำนวน 1000 ครั้ง แต่เนื่องจากความสามารถของเครื่องคอมพิวเตอร์ไม่สามารถทำงานได้กับข้อมูลขนาดใหญ่ จึงแบ่งสถานการณ์ 144 สถานการณ์ออกเป็น 9 ส่วนส่วนละ 16 สถานการณ์ โดยกำหนดให้ข้อมูลจากแฟ้ม sit1, sit2, … , sit9 เป็นข้อมูลนำเข้า และกำหนดเมทริกซ์ x11 และ x21 แทนข้อมูล x_1 และ x_2 ที่ได้จากสุ่มภายใต้สถานการณ์ในส่วนที่ 1 และในทำนองเดียวกันกับสถานการณ์ในแต่ละส่วน ดังนั้นเมทริกซ์ x19 และ x29 แทนข้อมูล x_1 และ x_2 ที่ได้จากสุ่มภายใต้สถานการณ์ในส่วนที่ 9 ต่อไปจะแสดงแผนผังแสดงขั้นตอนการสุ่มตัวอย่าง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3. จำลองค่าสัดส่วนการปฏิเสธสมมติฐาน H_0 เมื่อ H_0 เป็นจริง และ H_0 เป็นเท็จ สำหรับการทดสอบสมมูลภายใต้แนวคิดแบบคลาสสิก

ในการจำลองค่าสัดส่วน**การปฏิเสธสมมติฐาน H_0 เมื่อ H_0 เป็นจริง และ H_0 เป็นเท็จ**จะทำการจำลองที่ระดับนัยสำคัญ(α) 0.05 โดยกำหนดให้เมทริกซ์ pcls1, pcls2, … , pcls9 เก็บค่าสัดส่วนดังกล่าวทั้งสิ้น 144 สถานการณ์ ดังนั้นแต่ละเมทริกซ์จึงมีขนาด 16X1 โดยการจำลองค่าดังกล่าวทำได้โดยใช้ข้อมูลที่สุ่มได้จากขั้นตอนที่ 2. นั้นคือข้อมูลจากเมทริกซ์ x11, x21, … , x19, x29 และค่าพารามิเตอร์ $\theta_1, \theta_2, \Delta, n_1$ และ n_2 ที่กำหนดให้ทั้งหมด 144 สถานการณ์ เพื่อคำนวณค่า α^* และนำมาเปรียบเทียบกับค่า α เพื่อหาค่าสัดส่วนดังกล่าว โดยค่า α^* สามารถหาได้โดยวิธีประยุกต์ของพาเทิล-กุปตา (Modified Patel-Gupta test) ดังนี้ $\alpha^* = F(V/1+\gamma)$ โดยที่ F คือฟังก์ชันการแจกแจงสะสมของการแจกแจงแบบ F ด้วยองศาอิสระ $(1+\gamma^2)/(1+2\gamma)$ และ $n_1 + n_2 - 2$ เมื่อกำหนดให้

$$V = \left\{n_1(\hat{\theta}_1 - \hat{\tau})^2 + n_2(\hat{\theta}_2 - \hat{\tau})^2\right\}/s^2$$

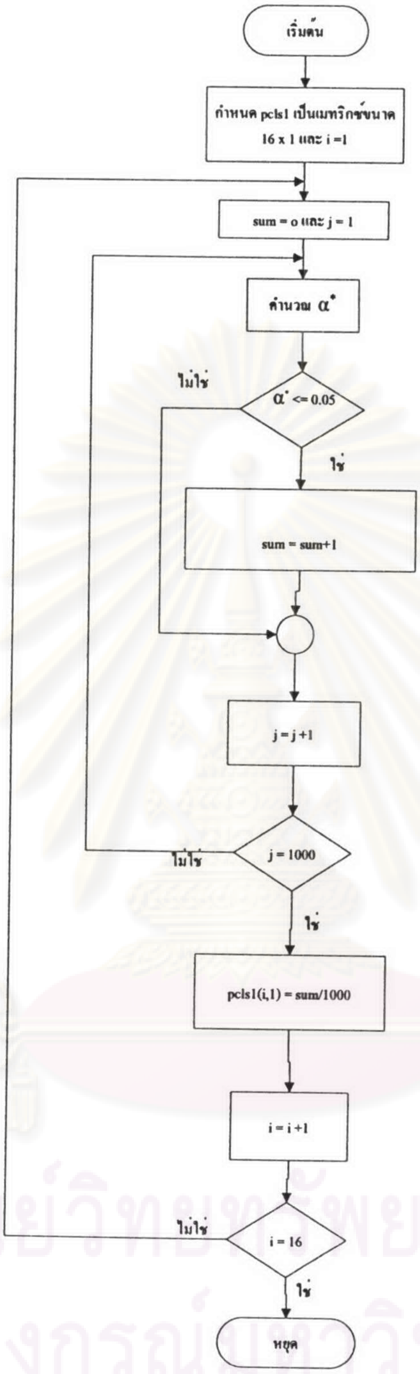
$$\hat{\tau} = (n_1\hat{\theta}_1 + n_2\hat{\theta}_2)/(n_1 + n_2)$$

$$s^2 = \left\{n_1\hat{\theta}_1(1-\hat{\theta}_1) + n_2\hat{\theta}_2(1-\hat{\theta}_2)\right\}/(n_1 + n_2)$$

$$\gamma = \left\{n_1 n_2/(n_1 + n_2)\right\}\left\{\Delta^2/s^2\right\}$$

และแผนผังแสดงขั้นตอนการจำลองค่าสัดส่วนเป็นดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

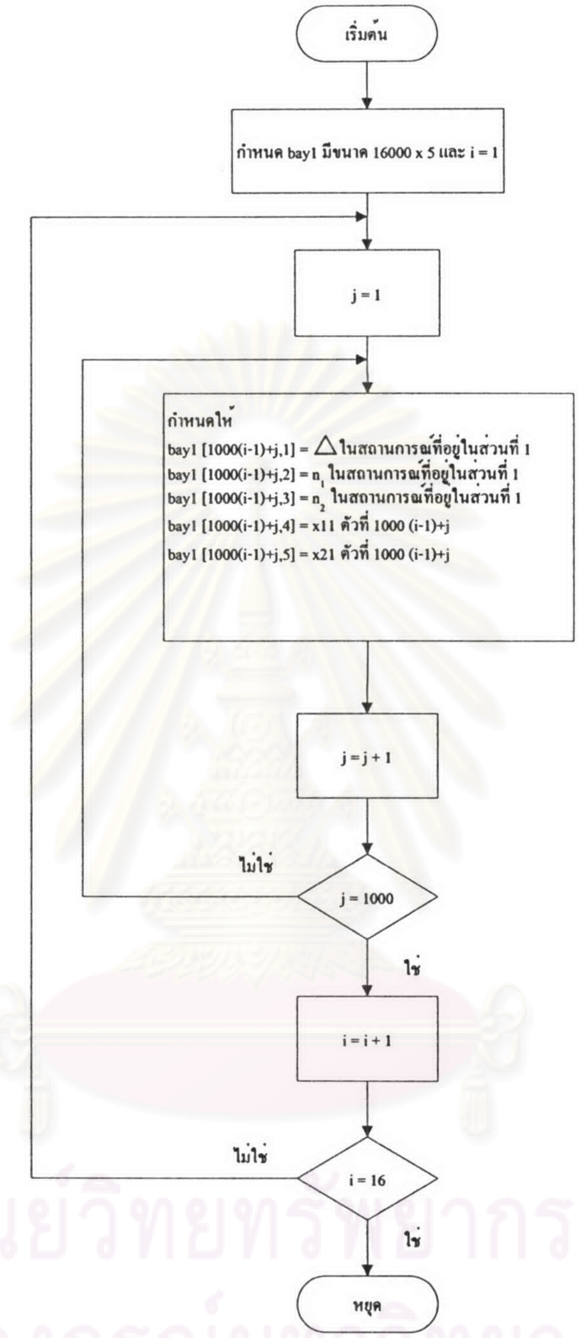


ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

4. เตรียมข้อมูลสำหรับการทดสอบสมมูลภายใต้แนวคิดแบบเบส์

เนื่องจากการทดสอบความสมมูลภายใต้แนวคิดแบบเบส์นั้นจำเป็นต้องอินทิเกรตสองชั้นเพื่อหาความน่าจะเป็นของฟังก์ชันการแจกแจงความน่าจะเป็นร่วม ของการแจกแจงเบตา (bivariate beta) ซึ่งการอินทิเกรตดังกล่าวไม่สามารถทำได้ด้วยวิธีการทางแคลคูลัส ดังนั้นจึงมีความจำเป็นต้องนำโปรแกรมทางคณิตศาสตร์เพื่อช่วยในการหาค่าอินทิเกรตดังกล่าว ในที่นี้ผู้วิจัยเลือกใช้โปรแกรม Mathcad เนื่องจากเป็นโปรแกรมที่สามารถคำนวณค่าอินทิเกรตดังกล่าวได้ แต่โปรแกรมดังกล่าวนั้นไม่เอื้อต่อการเขียนโปรแกรม ดังนั้นการใช้โปรแกรมดังกล่าวจึงจำเป็นต้องทำการจัดเรียงข้อมูล และพารามิเตอร์ต่างๆที่มีอยู่แล้วให้อยู่ในรูปเมทริกซ์ซึ่งแต่ละแถวของเมทริกซ์แทนด้วย Δ, n_1, n_2, x_1 และ x_2 ตามลำดับ โดยกำหนดชื่อเมทริกซ์ดังกล่าวคือ bay1, bay2, … , bay9 ซึ่งแต่ละเมทริกซ์แทนสถานการณ์ที่ต่างๆ 16 สถานการณ์มีขนาด 16,000X5 และแผนผังแสดงขั้นตอนเป็นดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5. จำลองค่าสัดส่วนการปฏิเสธสมมติฐาน H_0 เมื่อ H_0 เป็นจริงและ H_0 เป็นเท็จ สำหรับการทดสอบสมมูลภายใต้แนวคิดแบบเบส์

5.1 หาค่า $P(\theta \in \Theta_0^c|X)$

จาก $$P(\theta \in \Theta_0^c|X) = 1 - \int_{\Delta}^{1}\int_{0}^{\theta_2-\Delta} f(\theta_1,\theta_2)d\theta_1 d\theta_2 - \int_{0}^{1-\Delta}\int_{\theta_2+\Delta}^{1} f(\theta_1,\theta_2)d\theta_1 d\theta_2$$

ให้ $$P(\theta \in \Theta_0^c|X) = p(\Delta, n_1, n_2, \alpha_1, \beta_1, \alpha_2, \beta_2, x_1, x_2)$$

เมื่อ $$f(\theta_1,\theta_2) = \frac{\theta_1^{x_1+\alpha_1-1}(1-\theta_1)^{n_1-x_1+\beta_1-1}}{B(x_1+\alpha_1-1, n_1-x_1+\beta_1)} \cdot \frac{\theta_2^{x_2+\alpha_2-1}(1-\theta_2)^{n_2-x_2+\beta_2-1}}{B(x_2+\alpha_2-1, n_2-x_2+\beta_2)}$$

และ $$B(x_1+\alpha_1, n_1-x_1+\beta_1) = \frac{\Gamma(x_1+\alpha_1)\Gamma(n_1-x_1+\beta_1)}{\Gamma(x_1+\alpha_1+n_1-x_1+\beta_1)}$$

$$= \int_{0}^{1}\theta_1^{x_1+\alpha_1-1}(1-\theta_1)^{n_1-x_1+\beta_1-1}d\theta_1$$

และ $$B(x_2+\alpha_2, n_2-x_2+\beta_2) = \frac{\Gamma(x_2+\alpha_2)\Gamma(n_2-x_2+\beta_2)}{\Gamma(x_2+\alpha_2+n_2-x_2+\beta_2)}$$

$$= \int_{0}^{1}\theta_2^{x_2+\alpha_2-1}(1-\theta_2)^{n_2-x_2+\beta_2-1}d\theta_2$$

ให้ $$k(\Delta, n_1, n_2, \alpha_1, \beta_1, \alpha_2, \beta_2, x_1, x_2) = f(\Delta, n_1, n_2, \alpha_1, \beta_1, \alpha_2, \beta_2, x_1, x_2)$$

$$+ g(\Delta, n_1, n_2, \alpha_1, \beta_1, \alpha_2, \beta_2, x_1, x_2)$$

เมื่อ $$f(\Delta, n_1, n_2, \alpha_1, \beta_1, \alpha_2, \beta_2, x_1, x_2) = \int_{\Delta}^{1}\int_{0}^{\theta_2-\Delta}\theta_1^{x_1+\alpha_1-1}(1-\theta_1)^{n_1-x_1+\beta_1-1}$$

$$\theta_2^{x_2+\alpha_2-1}(1-\theta_2)^{n_2-x_2+\beta_2-1}d\theta_1 d\theta_2$$

และ $$g(\Delta, n_1, n_2, \alpha_1, \beta_1, \alpha_2, \beta_2, x_1, x_2) = \int_{0}^{1-\Delta}\int_{\theta_2+\Delta}^{1}\theta_1^{x_1+\alpha_1-1}(1-\theta_1)^{n_1-x_1+\beta_1-1}$$

$$\theta_2^{x_2+\alpha_2-1}(1-\theta_2)^{n_2-x_2+\beta_2-1}d\theta_1 d\theta_2$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

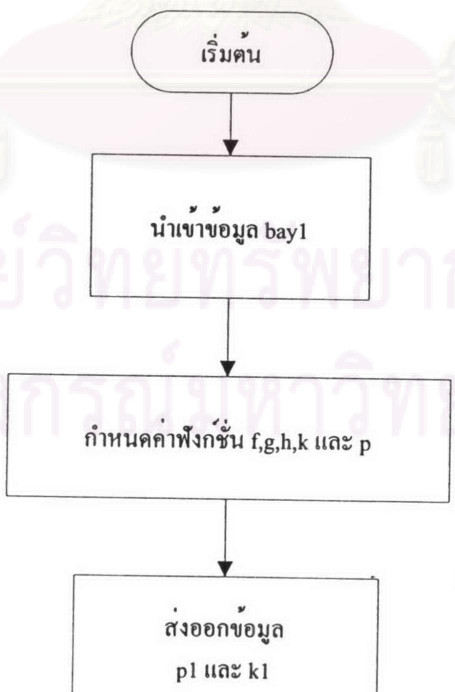
และให้ $h(\Delta,n_1,n_2,\alpha_1,\beta_1,\alpha_2,\beta_2,x_1,x_2) = \int_0^1 \theta_1^{x_1+\alpha_1-1}(1-\theta_1)^{n_1-x_1+\beta_1-1}d\theta_1$

$$\cdot \int_0^1 \theta_2^{x_2+\alpha_2-1}(1-\theta_2)^{n_2-x_2+\beta_2-1}d\theta_2$$

จะได้ $P(\theta \in \Theta_0^c|X) = p(\Delta,n_1,n_2,\alpha_1,\beta_1,\alpha_2,\beta_2,x_1,x_2)$

$$= 1 - \frac{k(\Delta,n_1,n_2,\alpha_1,\beta_1,\alpha_2,\beta_2,x_1,x_2)}{h(\Delta,n_1,n_2,\alpha_1,\beta_1,\alpha_2,\beta_2,x_1,x_2)}$$

การอินทิเกรตข้างต้นใช้โปรแกรมสำเร็จรูปทางคณิตศาสตร์ Mathcad และมีข้อมูลนำเข้าคือเมทริกซ์ bay1, … , bay9 ซึ่งได้จากการประมวลผลด้วยโปรแกรม S-PLUS ซึ่งเมทริกซ์ดังกล่าวยังไม่สามารถนำเข้าไปประมวลผลในโปรแกรม Mathcad ได้เนื่องจากเมทริกซ์ที่ได้จากการประมวลผลโดยโปรแกรม S-PLUS จะเกิดแถวแรกซึ่งเป็นแถวว่างเสมอ แต่การประมวลผลของโปรแกรม Mathcad จะเริ่มอ่านข้อมูลตั้งแต่แถวแรก ดังนั้นก่อนการประมวลผลด้วยโปรแกรม Mathcad จึงต้องลบแถวแรกซึ่งเป็นแถวว่างของเมทริกซ์ bay1, … , bay9 ออกเสียก่อนจึงสามารถประมวลผลได้ สำหรับค่า $P(\theta \in \Theta_0^c|X)$ ที่ได้จะเก็บไว้ในเมทริกซ์ใช้ชื่อว่า p1, … , p9 และค่าฟังก์ชัน k ที่ได้จะเก็บไว้ในเมทริกซ์ใช้ชื่อว่า k1, … , k9 เช่นเดียวกัน สำหรับแผนผังแสดงขั้นตอนเป็นดังนี้



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5.2 หาค่า $P(\theta \in \Theta_0^c | X)$ ด้วยการประมาณด้วยการแจกแจงปกติ

เนื่องจากโปรแกรม Mathcad นั้นสามารถเก็บค่าข้อมูลตัวเลขที่มีค่าเข้าใกล้ 0 ได้มากๆไม่ต่ำกว่า 10^{-307} แต่เนื่องจากค่าที่ได้จากฟังก์ชัน k ในขั้นตอน 5.1 นั้นในบางสถานการณ์มีค่าต่ำกว่า 10^{-307} ทำให้โปรแกรมเก็บค่าข้อมูลดังกล่าวเป็น 0 ดังนั้นค่า $P(\theta \in \Theta_0^c | X)$ จะไม่สามารถหาค่าได้ในบางกรณี ทำให้ต้องหา $P(\theta \in \Theta_0^c | X)$ ด้วยวิธีการประมาณการแจกแจงเบตาด้วยการแจกแจงปกติ ซึ่งสามารถทำได้ดังนี้

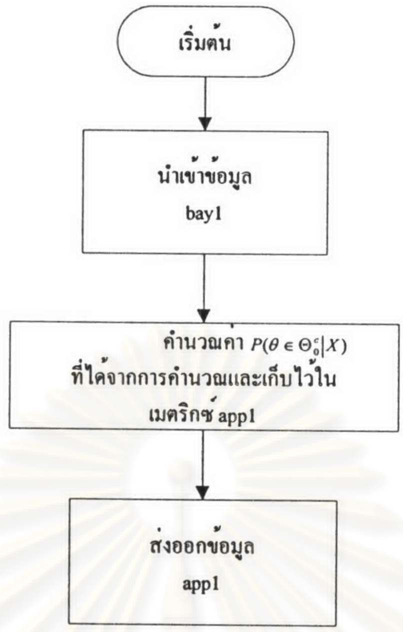
กำหนดให้ $X \sim Beta(\alpha, \beta)$ เมื่อ $\alpha = \beta \to \infty$ แล้ว X จะมีการแจกเข้าใกล้การแจกแจงปกติที่มีค่าเฉลี่ย μ และค่าแปรปรวน σ^2 โดยที่ $\mu = E(X) = \dfrac{\alpha}{\alpha + \beta}$ และ

$$\sigma^2 = Var(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

ซึ่งในการประมาณนั้นสามารถทำได้ดีเนื่องจากกรณีที่ทำให้ค่าที่ได้จากฟังก์ชัน k ที่มีค่าต่ำมากๆนั้นเกิดจากค่าพารามิเตอร์ที่มีค่าใกล้เคียงกันและมีค่ามาก ดังนั้นจึงสามารถใช้การแจกแจงปกติประมาณการแจกแจงเบตาได้ดังนี้

เนื่องจาก $f(\theta_1, \theta_2) = \pi(\theta_1 | x_1) \cdot \pi(\theta_2 | x_2)$ และ θ_1 และ θ_2 มีการแจกแจงเหมือนกัน นั่นคือ $Beta(x_1 + \alpha_1, n_1 - x_1 + \beta_1)$ และ $Beta(x_2 + \alpha_2, n_2 - x_2 + \beta_2)$ ตามลำดับ เมื่อประมาณการแจกแจงเบตาด้วยการแจกแจงปกติแล้ว จะได้ θ_i มีการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับ $\dfrac{x_i + \alpha_i}{x_i + \alpha_i + n_i - x_i + \beta_i}$ และค่าแปรปรวนเท่ากับ $\dfrac{(x_i + \alpha_i)(n_i - x_i + \beta_i)}{(x_i + \alpha_i + n_i - x_i + \beta_i)^2(x_i + \alpha_i + n_i - x_i + \beta_i + 1_i)}$ เมื่อ $i = 1,2$ และสำหรับค่า $P(\theta \in \Theta_0^c | X)$ ที่ประมาณได้จะเก็บไว้ในเมทริกซ์ใช้ชื่อว่า app1 ถึง app9 สำหรับแผนผังแสดงขั้นตอนเป็นดังนี้

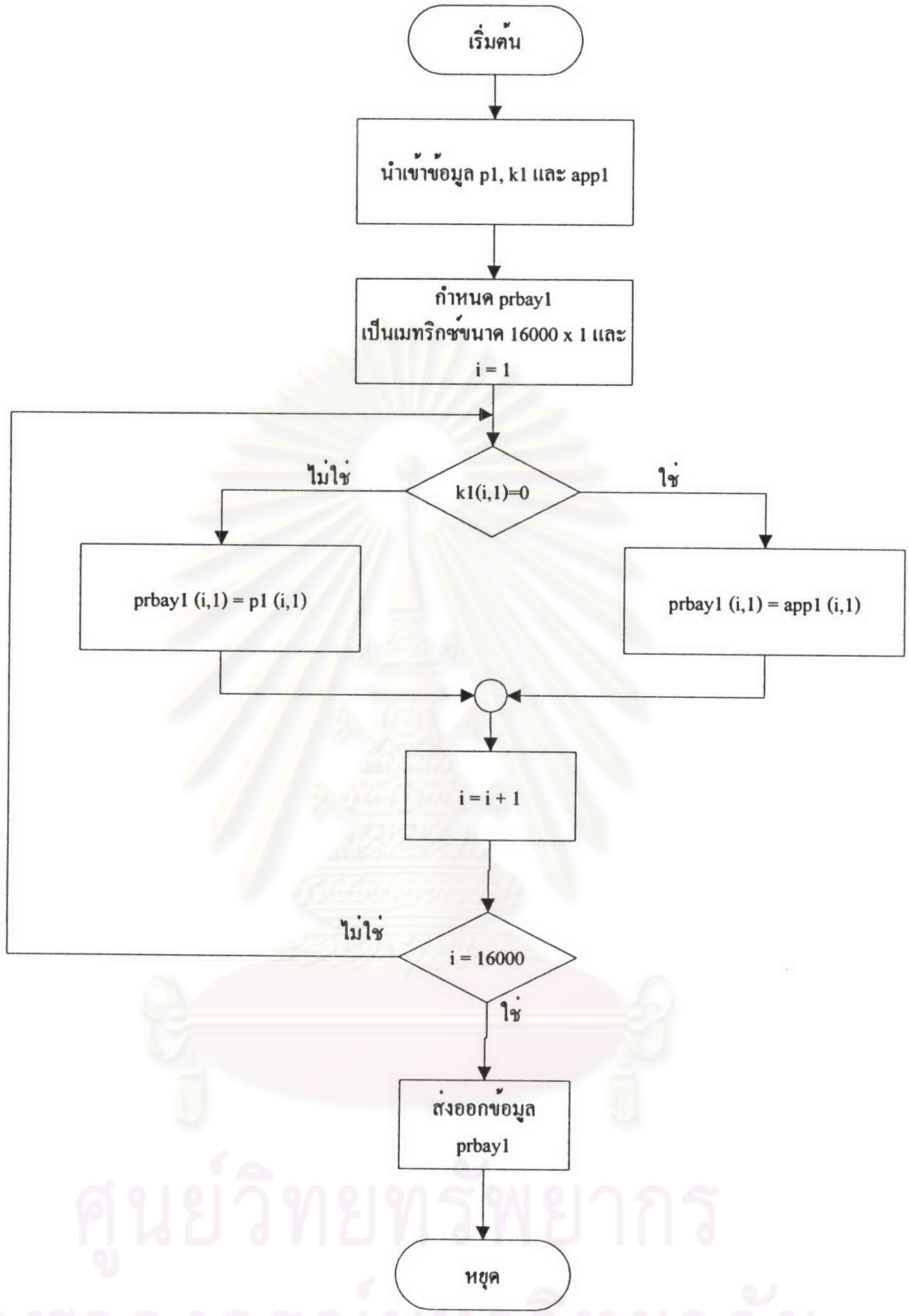
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



5.3 หาค่า $P(\theta \in \Theta_0^c|X)$ ที่เหมาะสม

เนื่องจากค่า $P(\theta \in \Theta_0^c|X)$ ที่ได้จากขั้นตอนที่ 5.1 นั้นไม่สามารถหาค่าได้ทุกค่า ดังนั้นค่าใดที่ไม่สามารถหาได้ จะเลือกใช้ค่าประมาณของ $P(\theta \in \Theta_0^c|X)$ ที่ได้จากขั้นตอนที่ 5.2 โดยกำหนดเมทริกซ์ prbay1, … , prbay9 เป็นเมทริกซ์ที่แต่ละเมทริกซ์มีขนาด 16,000X1 และแผนผังแสดงขั้นตอนเป็นดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

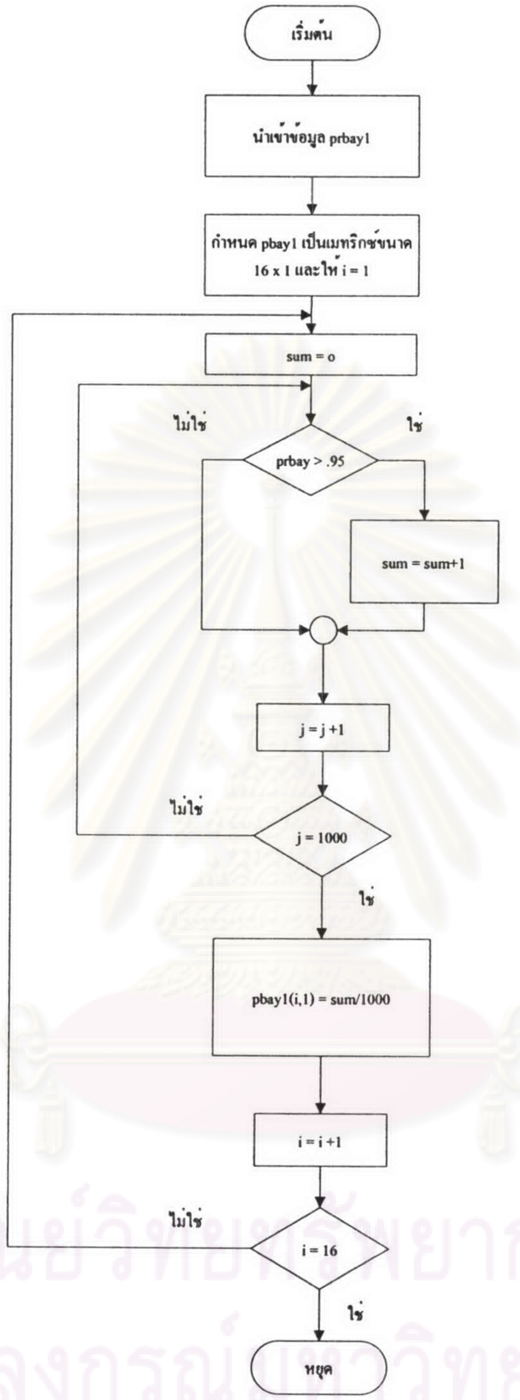


ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5.4 จำลองค่าสัดส่วนการปฏิเสธสมมติฐาน H_0

ในการจำลองค่าสัดส่วนดังกล่าว จะทำการจำลองค่าสัดส่วนการปฏิเสธ H_0 เมื่อ H_0 เป็นจริง และ H_0 เป็นเท็จ ด้วยสถานการณ์ที่แตกต่างกัน 144 สถานการณ์ เมื่อกำหนดค่า k เท่ากับ 0.95 โดยกำหนดให้เมทริกซ์ pbay1, pbay2, … , pbay9 เก็บค่าสัดส่วนดังกล่าวเมทริกซ์ละ 16 สถานการณ์ ดังนั้นแต่ละเมทริกซ์จึงมีขนาด 16X1 โดยพิจารณาจากค่า $P(\theta \in \Theta_0^c | X) > k$ หรือไม่ ถ้ามากกว่าจะตัดสินใจปฏิเสธสมมติฐาน และแผนผังแสดงขั้นตอนเป็นดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย


เริ่มต้น
นำเข้าข้อมูล prbay1
กำหนด pbay1 เป็นเมทริกซ์ขนาด 16 x 1 และให้ i = 1
sum = o
ไม่ใช่
ใช่
prbay > .95
sum = sum+1
j = j +1
ไม่ใช่
j = 1000
ใช่
pbay1(i,1) = sum/1000
i = i +1
ไม่ใช่
i = 16
ใช่
หยุด

6. สรุปผลการเปรียบเทียบระหว่างสองวิธีในสถานการณ์ต่างๆ

เมื่อได้ค่าสัดส่วนจำลองในการปฏิเสธสมมติฐาน H_0 เมื่อ H_0 เป็นจริง และค่าจำลองในการปฏิเสธสมมติฐาน H_0 เมื่อ H_0 เป็นเท็จของทั้งสองวิธีแล้ว นำผลที่ได้มาสรุปลงในตารางแล้วเปรียบเทียบผลที่ได้ในแต่ละสถานการณ์

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย